Multimodal fitting of atypical size distributions from AERONET

Michael Taylor
Stelios Kazadzis
Evangelos Gerasopoulos

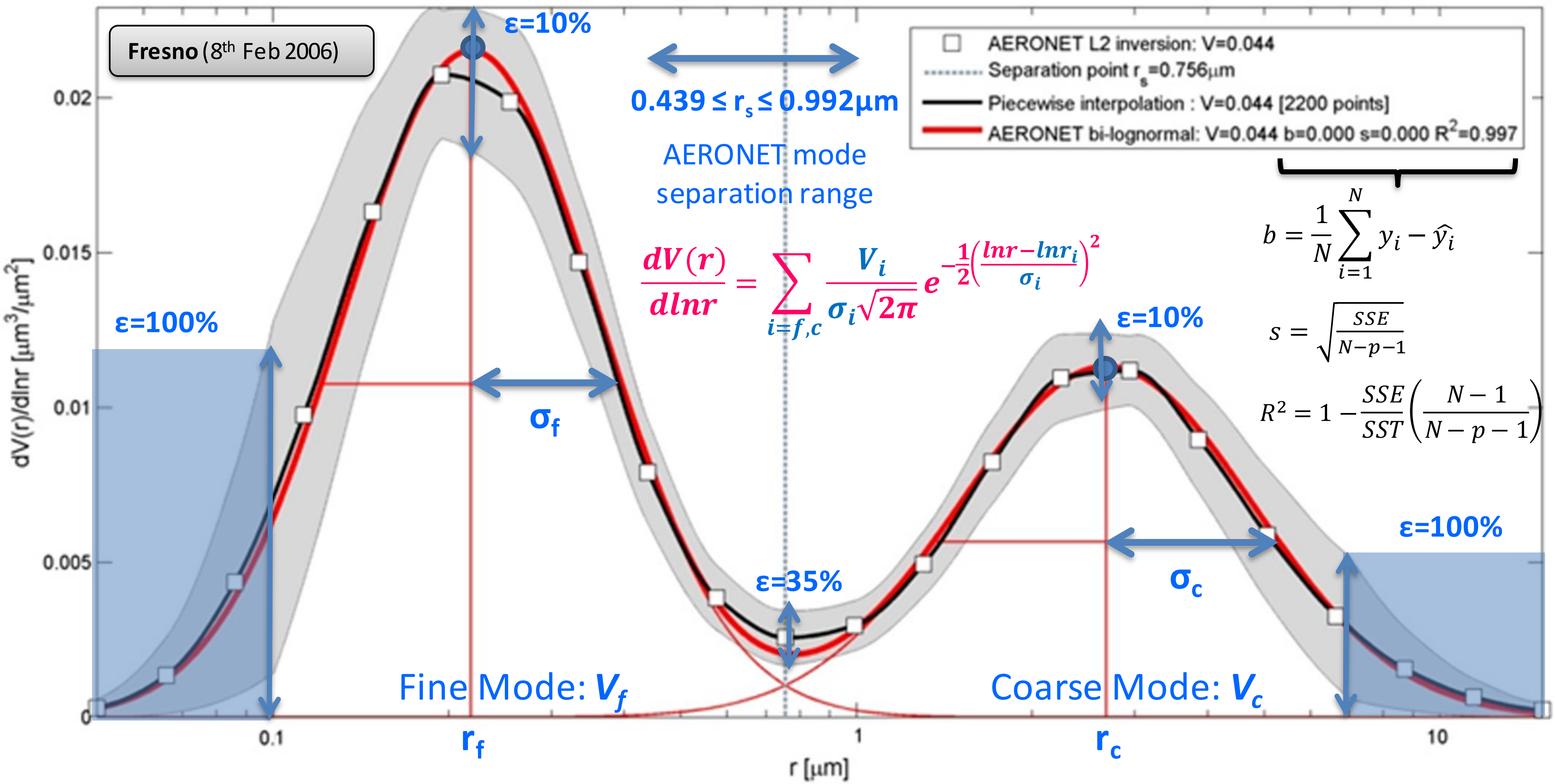
URL: *http://apcg.meteo.noa.gr*
eMail: *mtaylor@noa.gr*

OVERVIEW

1. *Typical & atypical distributions*
2. *Two new fitting methods*
3. *An interesting new case*
4. *Potential impact & a wish list ☺*

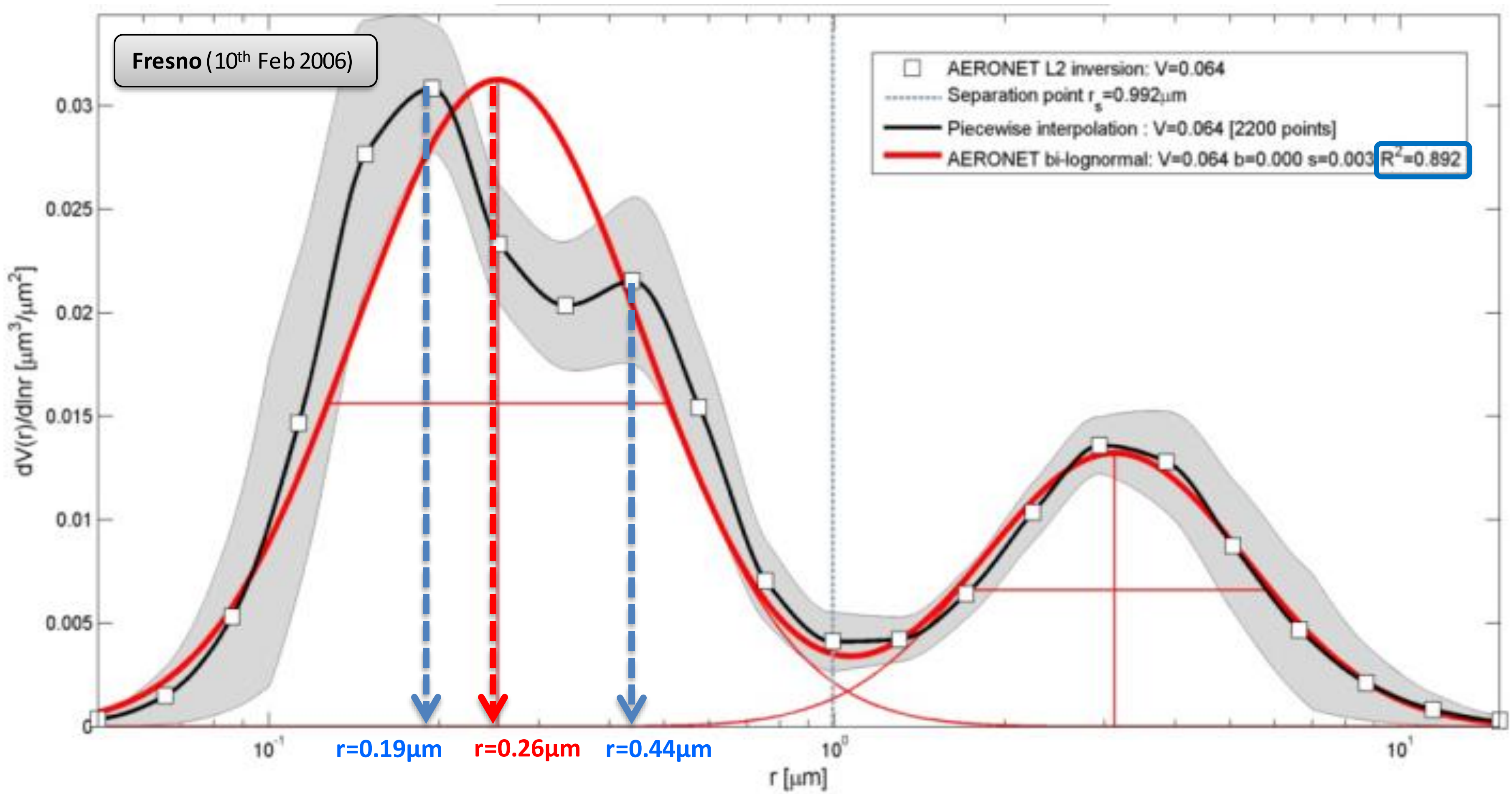
1. Typical & atypical distributions

1a) A typical (bi-modal) size distribution



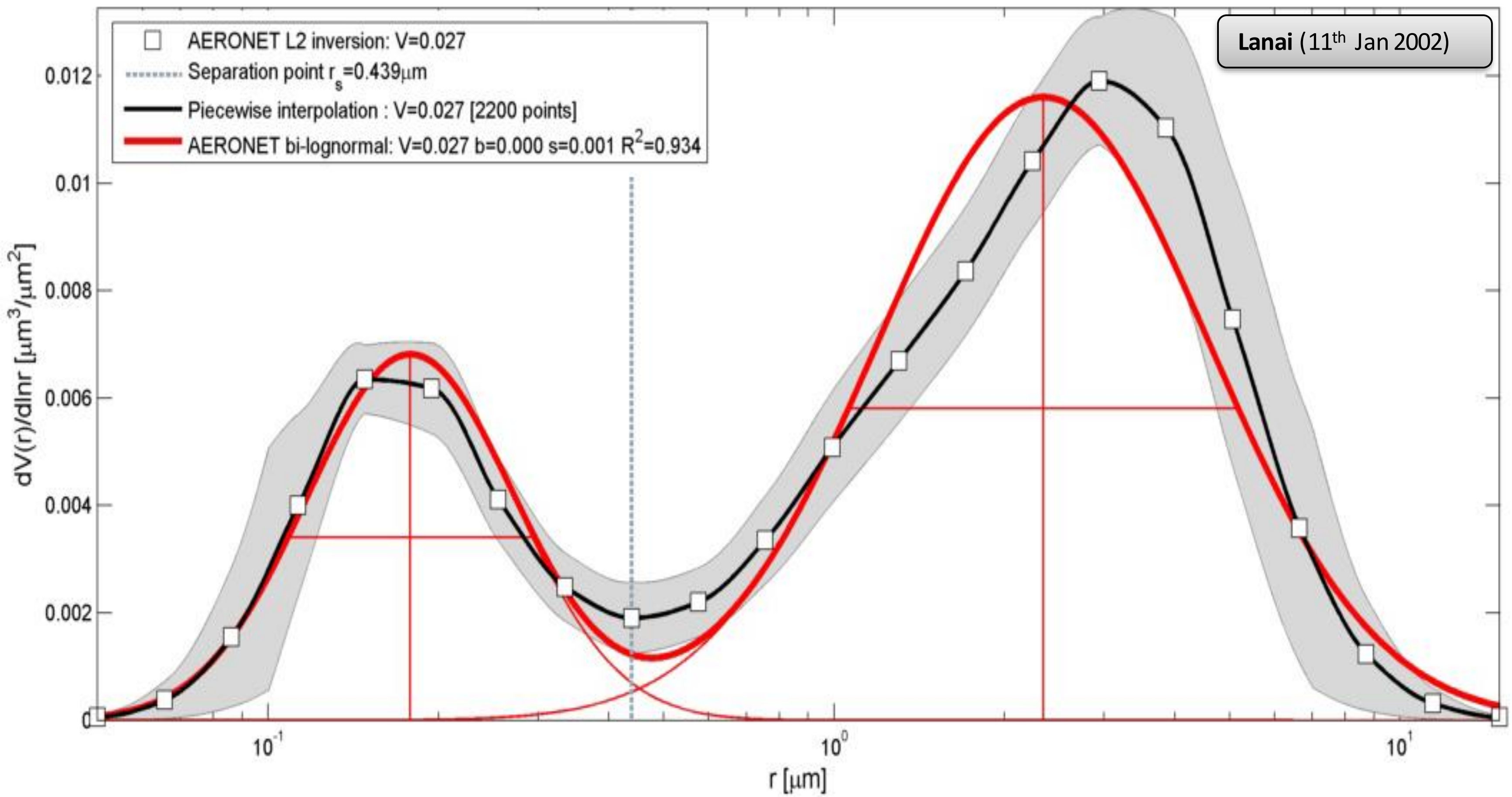
$$V_i = \int_{r_1}^{r_2} \frac{dV(r)}{dlnr}\, dlnr \qquad \ln r_i = \frac{\int_{r_1}^{r_2} \ln r \frac{dV(r)}{d\ln r}\, d\ln r}{\int_{r_1}^{r_2} \frac{dV(r)}{d\ln r}\, d\ln r} \qquad \sigma_i = \sqrt{\frac{\int_{r_1}^{r_2} (\ln r - \ln r_i)^2 \frac{dV(r)}{d\ln r}\, d\ln r}{\int_{r_1}^{r_2} \frac{dV(r)}{d\ln r}\, d\ln r}}$$

1a) 2 days later…

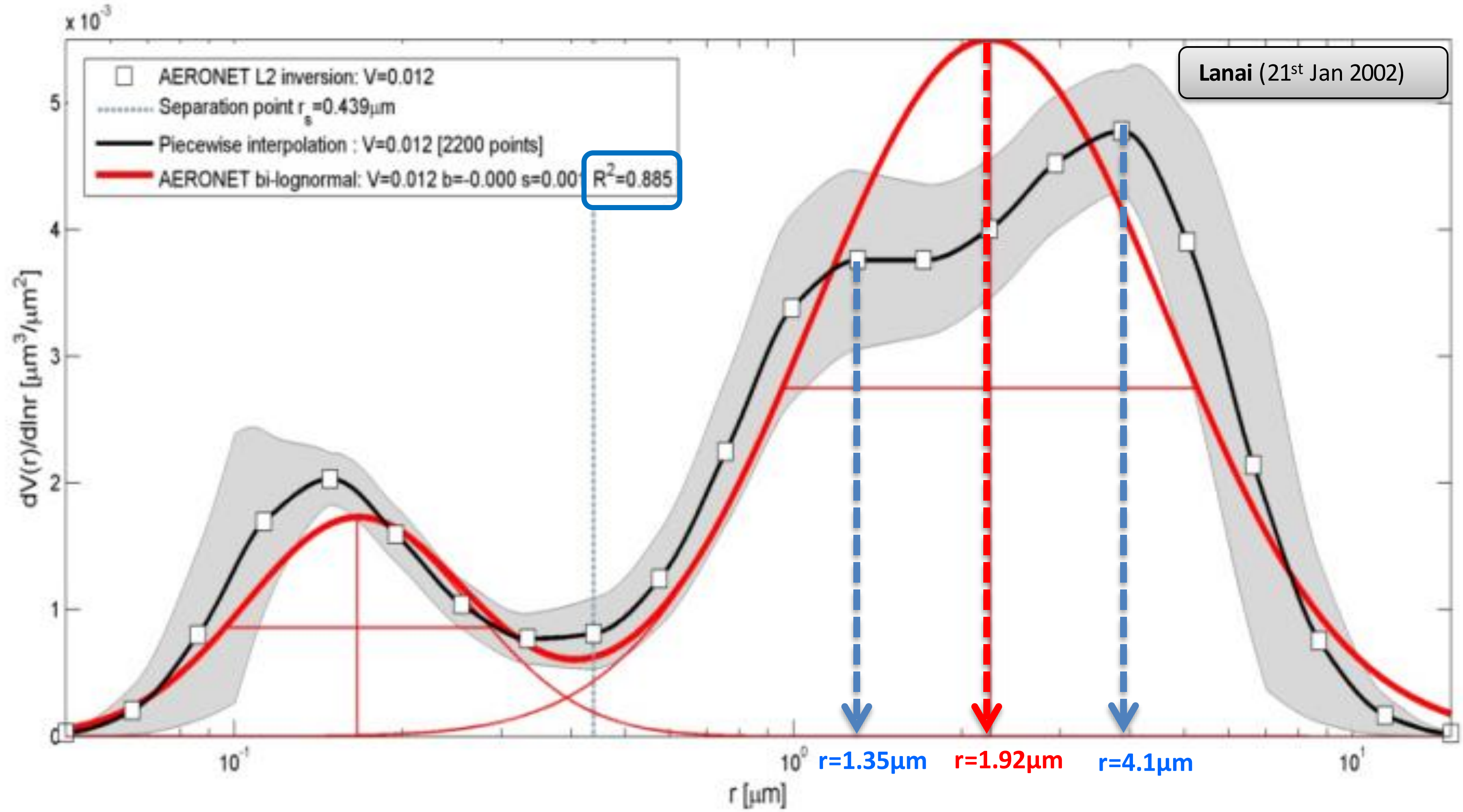


1. **mis-identification of "fine" modes**
2. **creation of a "ghost" (non-physical) fine mode**
3. **drop in goodness of fit:** R^2=0.997 → R^2=0.892

1a) another typical (bi-modal) size distribution



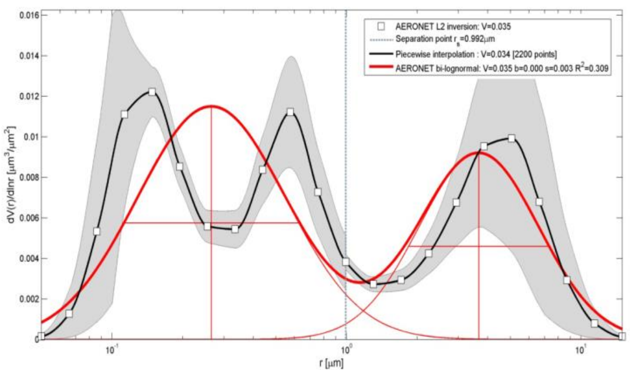
1a) 10 days later…



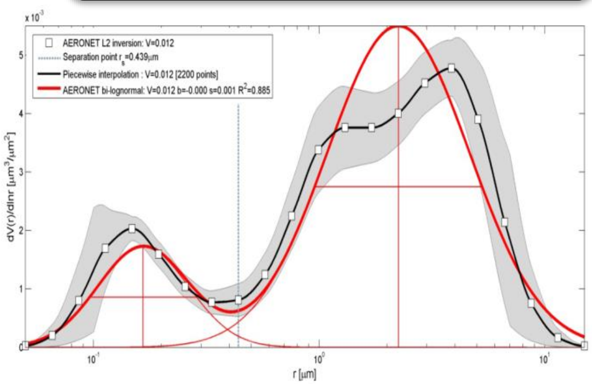
1. mis-identification of "coarse" mode(s)
2. creation of a "ghost" (non-physical) coarse mode
3. **drop in goodness of fit:** R^2=0.934 → R^2=0.885

1b) IDEA: an initial taxonomy of atypical distributions

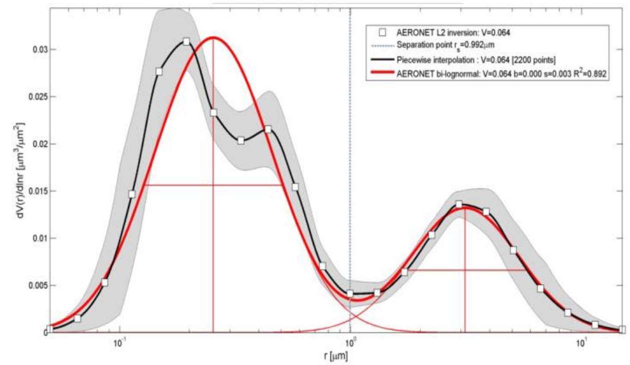
Washington-GSFC (23^{th} Jun 1993)
Triple peak (Pinatubo ash effect)



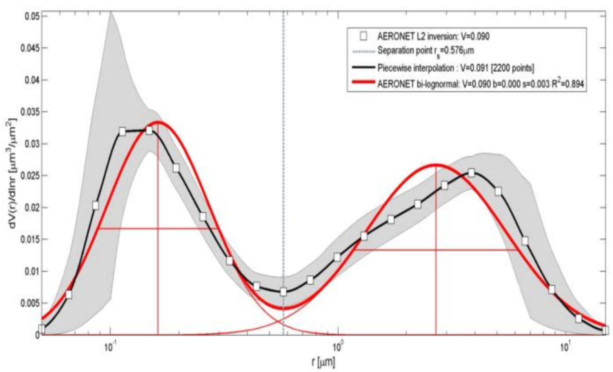
Lanai (21^{st} Jan 2002)
Double-coarse peak



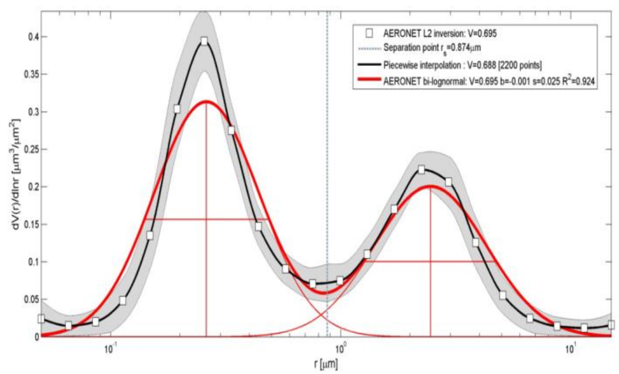
Fresno (10^{th} Feb 2006)
Double-fine peak



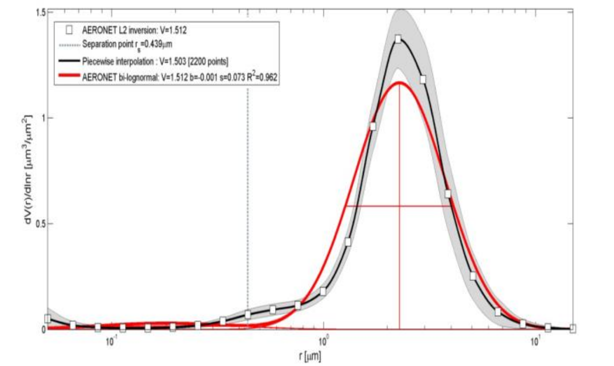
Beijing (18^{th} Feb 2011)
Skewed fine & coarse peaks



Beijing (23^{th} Feb 2011)
Elevated mid-point



Solar Village (29^{th} Mar 2011)
Quenched fine mode



Q. Anyone interested in collaborating to extend our *database* of atypical events ?

1c) IDEA: can we use R^2 to detect atypical events?

AERONET Site	Date	Observation	R^2
GSFC-Washington	23-Jun-93	Triple Peak	0.309
Lanai	20-Jan-02	Double Coarse Peak	0.833
Fresno	10-Feb-06	Double Fine Peak	0.892
Beijing	18-Feb-11	Skewed Fine & Coarse Peaks	0.894
Beijing	23-Feb-11	Elevated Mid-Point	0.924
Solar Village	29-Mar-11	Quenched Fine Peak	0.962
Frenso	**08-Feb-06**	**Bi-modal**	**0.997**

Strongly atypical

Moderately atypical

Weakly atypical

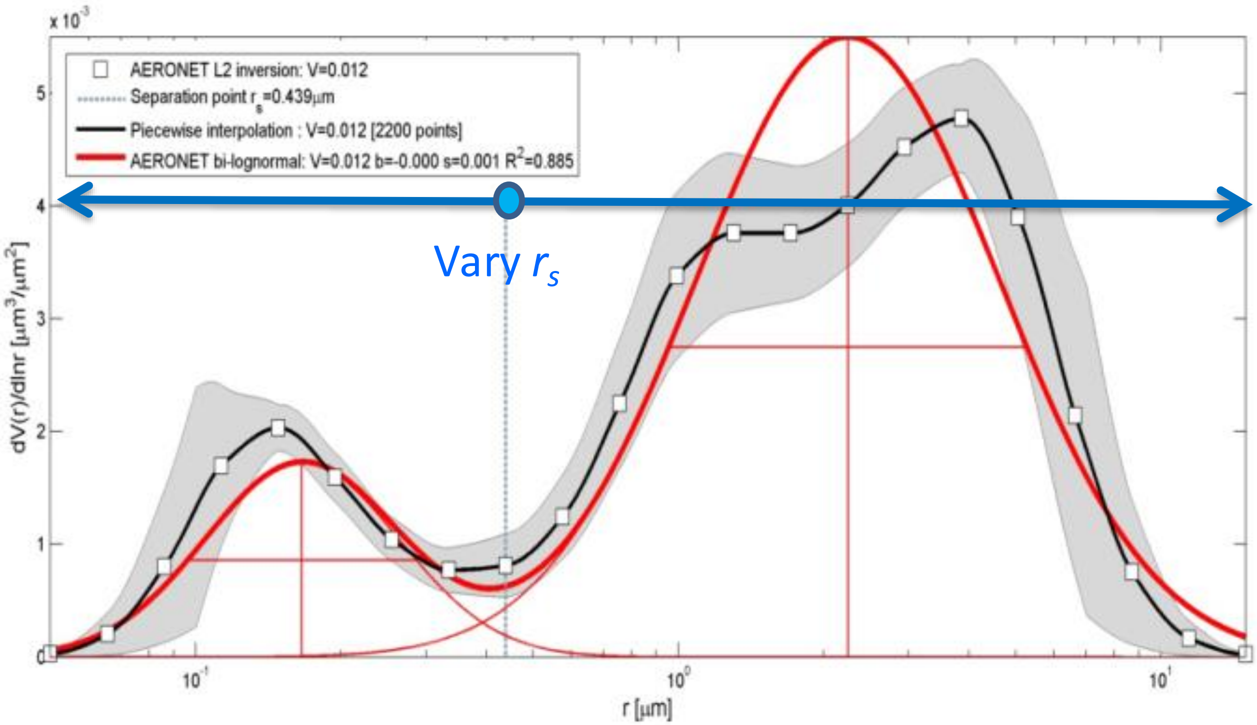
Typical case

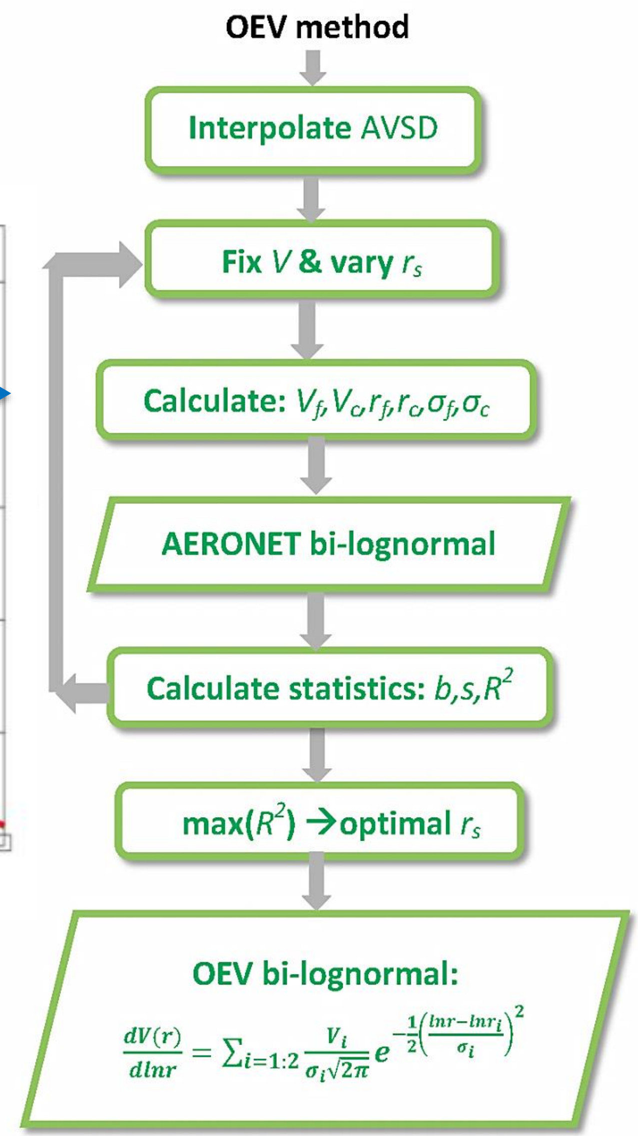
2. Two new fitting methods

Taylor, Kazadzis, Gerasopoulous (2014): AMT 7, 839-858

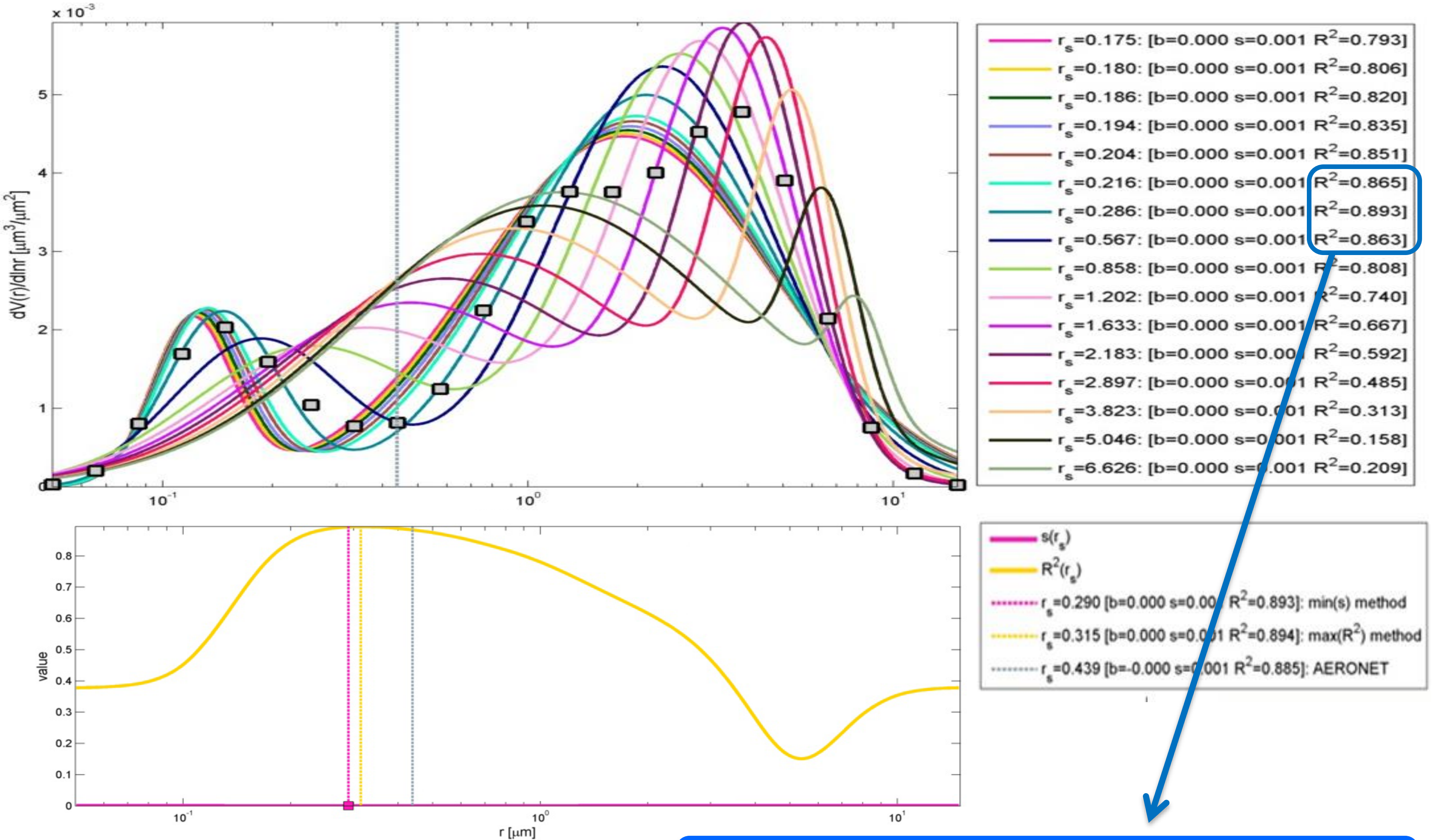
2a) Optimized Equivalent Volume (OEV) method

Lanai (21^{st} Jan 2002)





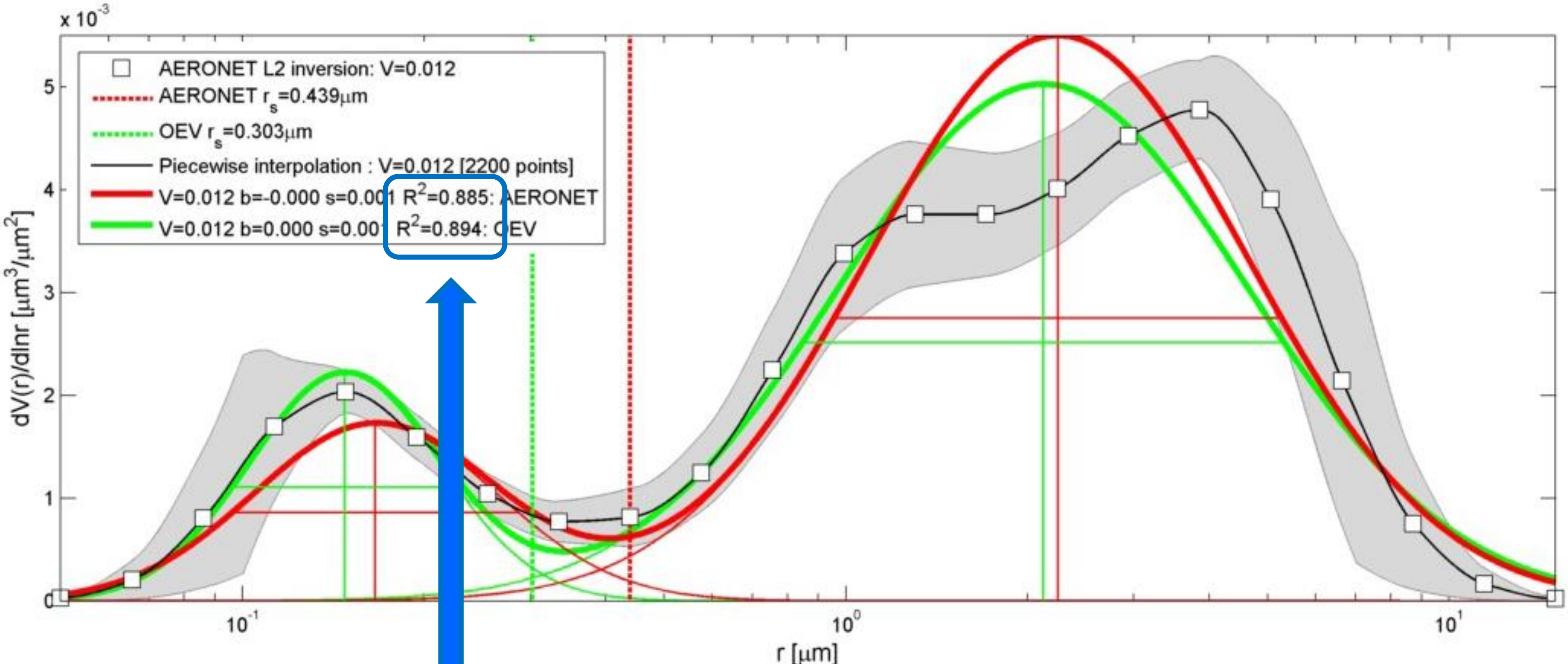
2a) OEV method: varying r_s



Max(R^2) → criterion for identifying optimal r_s

2a) OEV method: *comparison with AERONET bi-modal*

Lanai (21st Jan 2002)

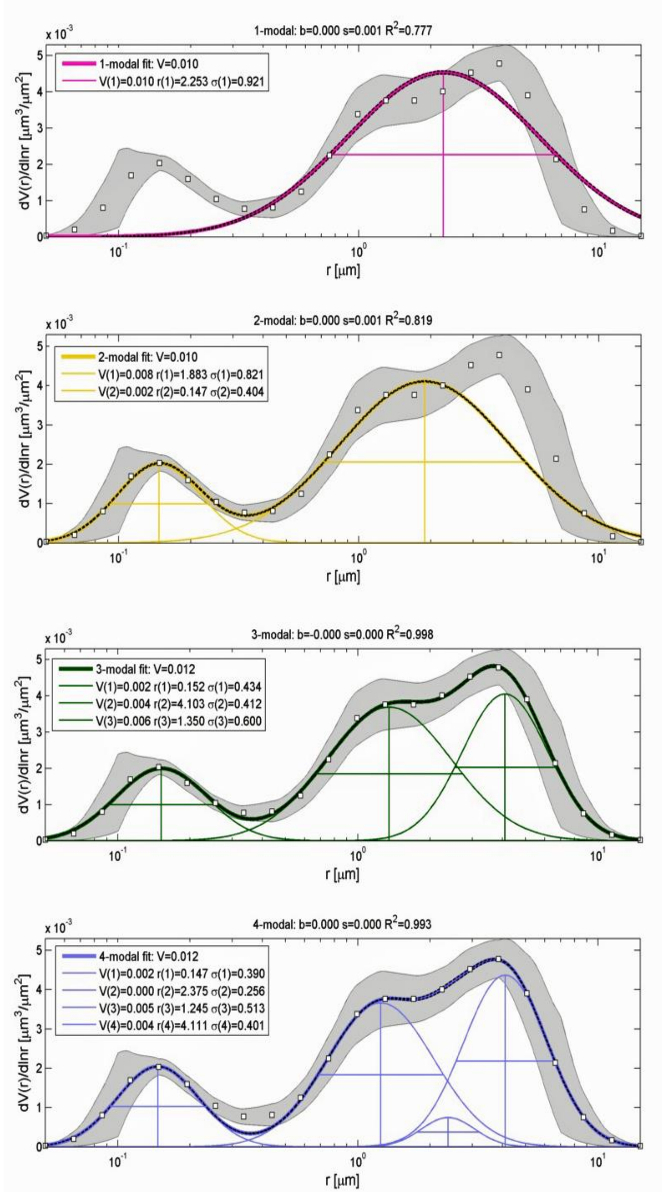


TINY quantitative but <u>not</u> **qualitative improvement**

2b) Gaussian Mixture Model (GMM) method *varying n-modes*

IERSD - NOA

Lanai (21st Jan 2002)



GMM method

- Interpolate AVSD
- Vary n-modes
- Calculate: $V(i=1:n), r(i=1:n), \sigma(i=1:n)$
- GMM *n*-lognormal
- Calculate statistics: b, s, R^2
- H_o: $R^2(n+1) = R^2(n)$ → optimal *n*
- GMM n-lognormal:

$$\frac{dV(r)}{dlnr} = \sum_{i=1:n} \frac{V_i}{\sigma_i\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{lnr - lnr_i}{\sigma_i}\right)^2}$$

2b) GMM method: *the stopping condition*

$$\rho \approx \sqrt{R^2}$$

Harel (2009): App Stat 36(10): 1109-1118

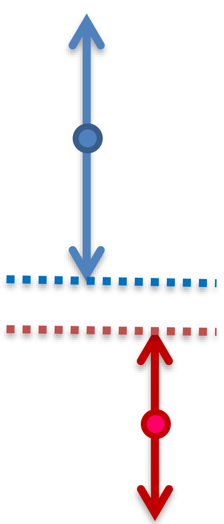
(maximum *RE*=0.060%)

$$F(\rho) = \frac{1}{2} ln \frac{1+\rho}{1-\rho}$$

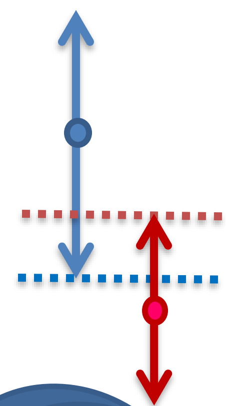
Fisher (1921): Metron 1: 3–32

$$CI = F(\rho) \pm 1.96/\sqrt{N-3}$$

(95% confidence level)



CASE 1: Two values of $F(\rho)$ (and hence R^2) show a significant statistical difference when the lower CI of the larger $F(\rho)$ value does not overlap the upper confidence limit of the smaller $F(\rho)$ value



CASE 2: In the event of an overlap, there is a statistical difference when t>1.96

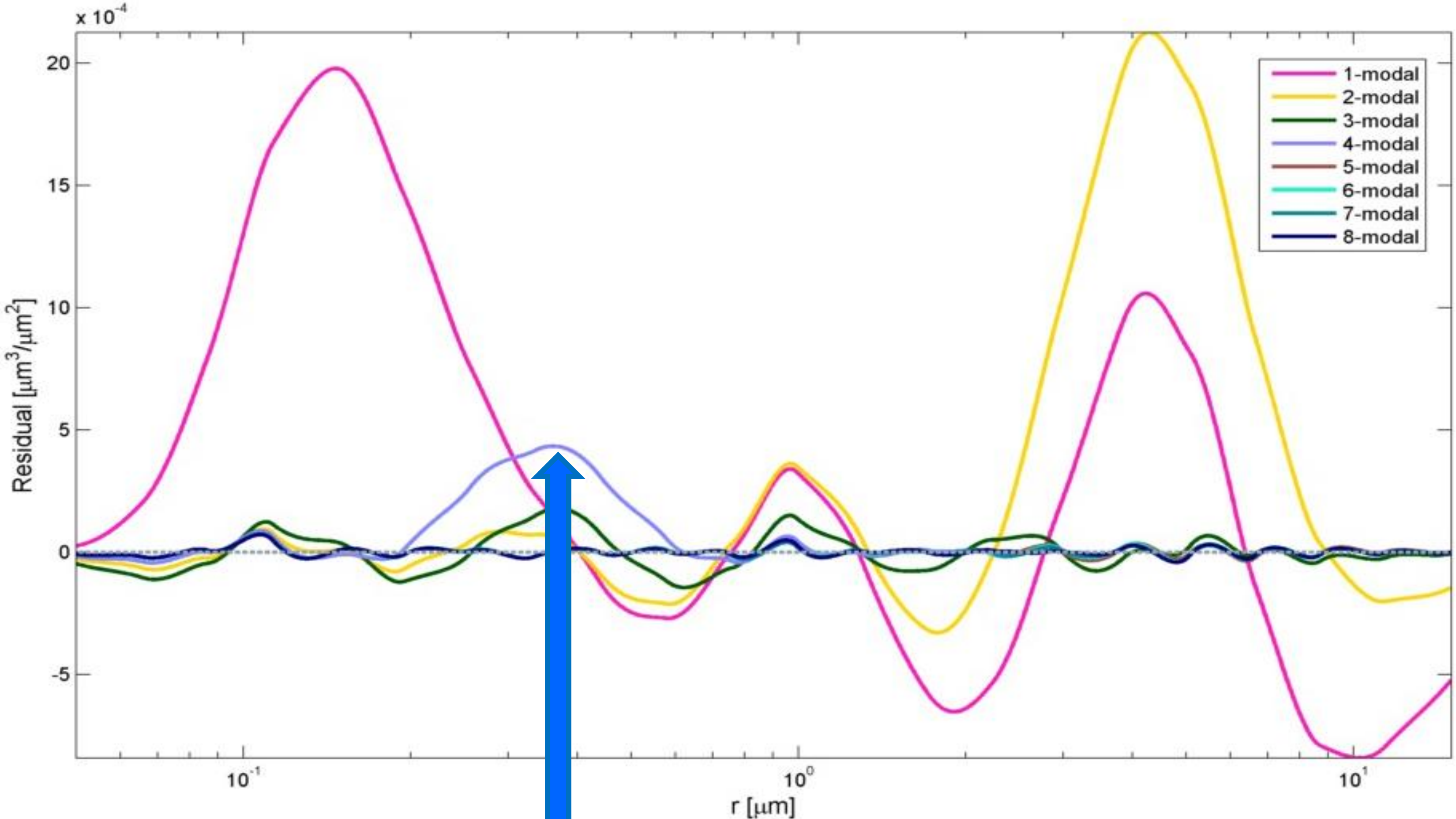
$$t = \left| \frac{F(\rho_1) - F(\rho_2)}{\sqrt{\frac{1}{N_1 - 3} + \frac{1}{N_2 - 3}}} \right|$$

Welch (1947): Biometrika 34 (1–2): 28–35

n Modes	$R^2(n)$	$R^2(n+1)$	$F(\rho_1)$	$F(\rho_2)$	$CI_1(l)$	$CI_1(u)$	$CI_2(l)$	$CI_2(u)$	t-Welch
1	0.777								
2	0.777	0.819	1.38	1.50	1.34	1.54	1.46	1.54	3.87
3	0.819	0.998	1.50	3.80	1.46	3.84	3.76	3.84	76.26
4	0.998	0.993	3.80	3.17	3.76	3.21	3.13	3.21	20.80

2b) GMM method: *residual analysis*

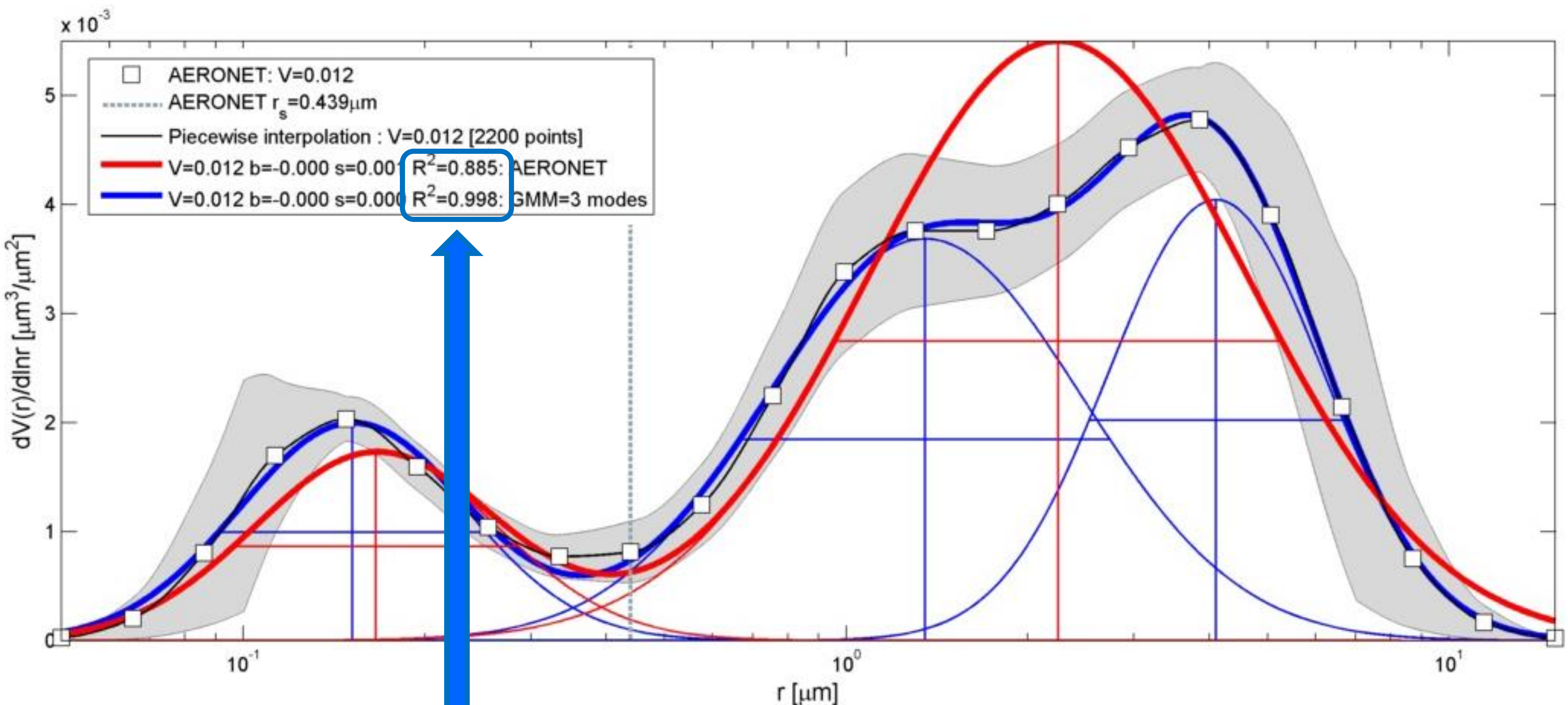
Lanai (21st Jan 2002)



LARGE **reduction in residuals when** *n* = 3 <u>and</u> **a small increase when** *n* = 4

2b) GMM method: *comparison with AERONET bi-modal*

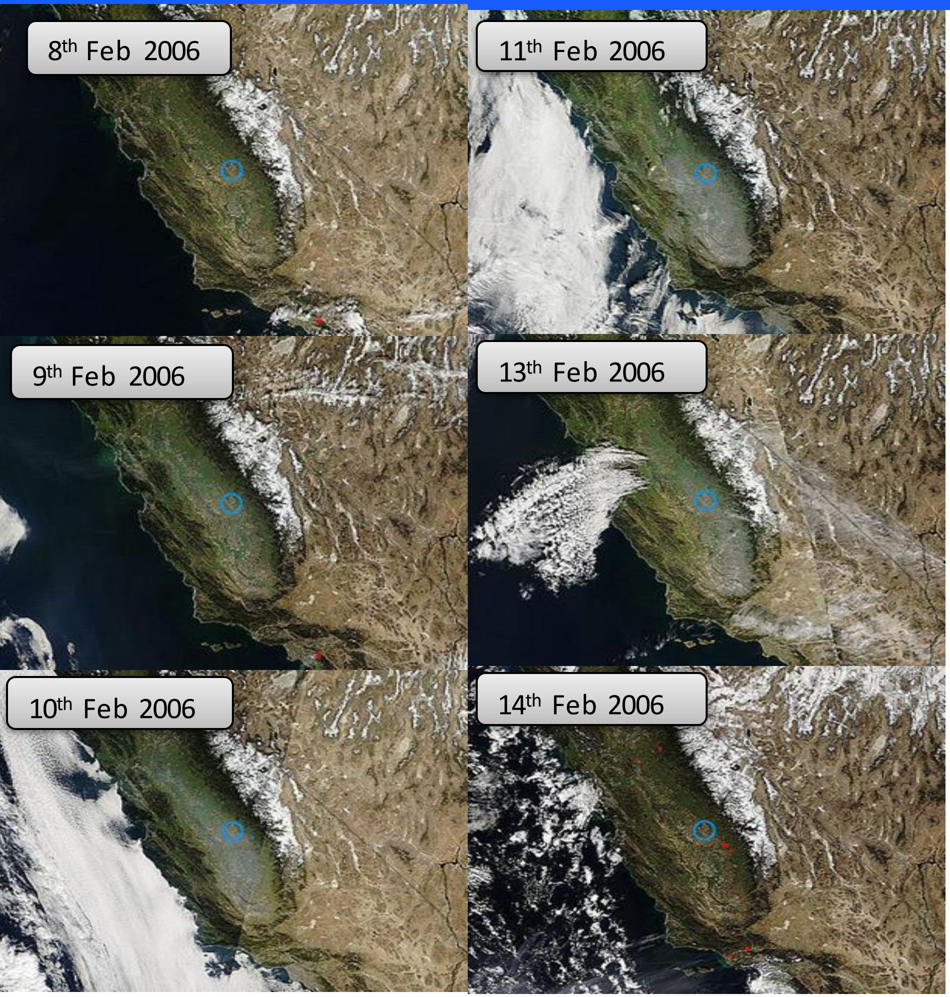
Lanai (21st Jan 2002)

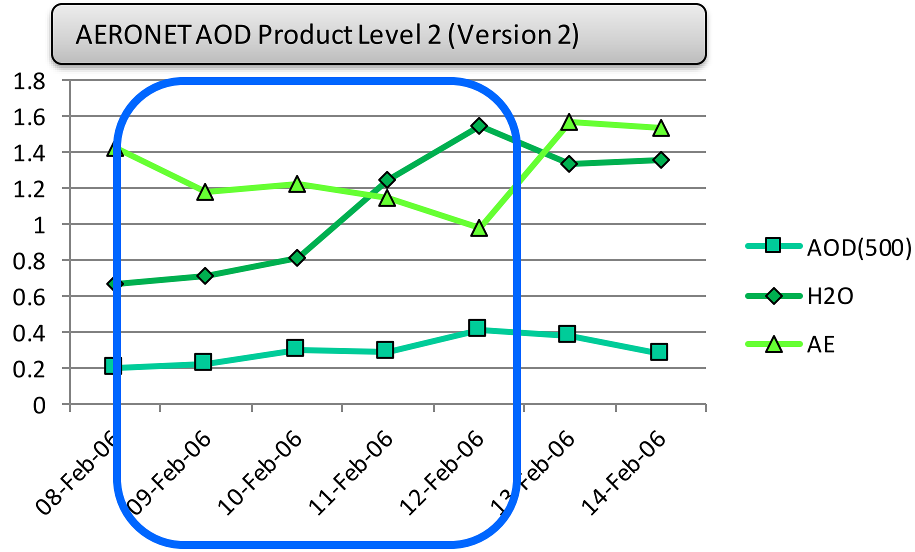


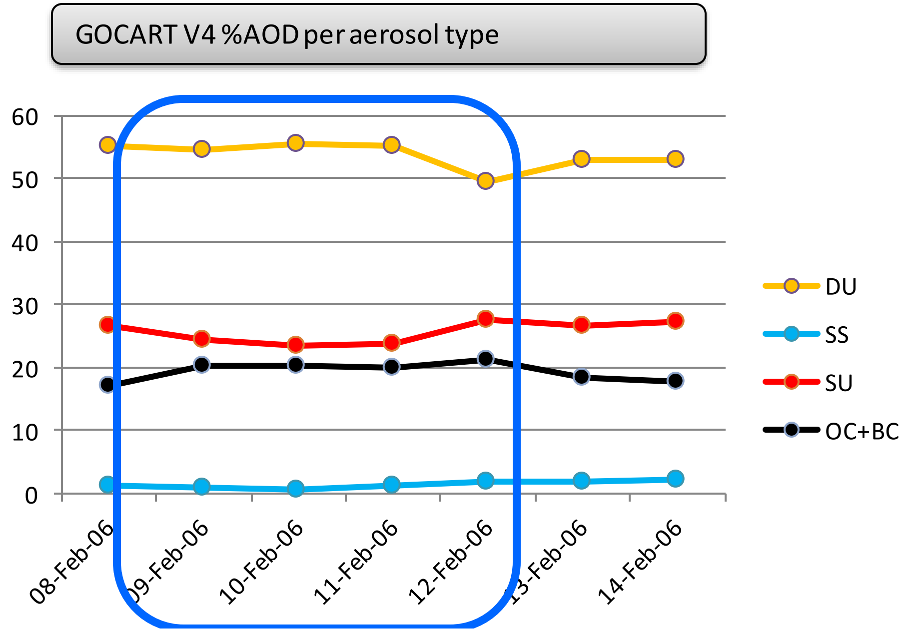
LARGE quantitative and qualitative improvement

3. An interesting new case

3a) Fresno: 8th–14th Feb 2006: *MODIS 2km + aerosol props.*

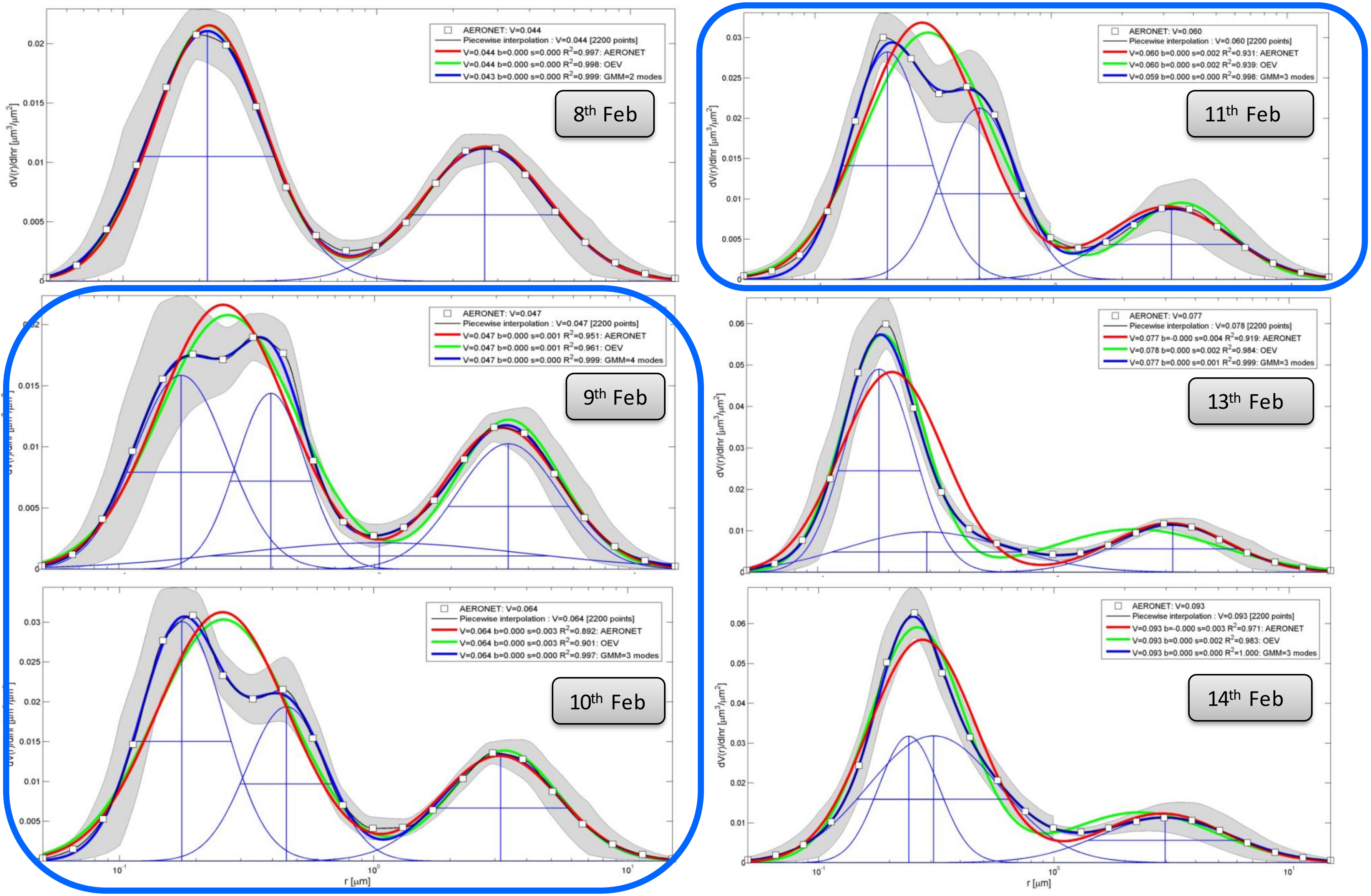






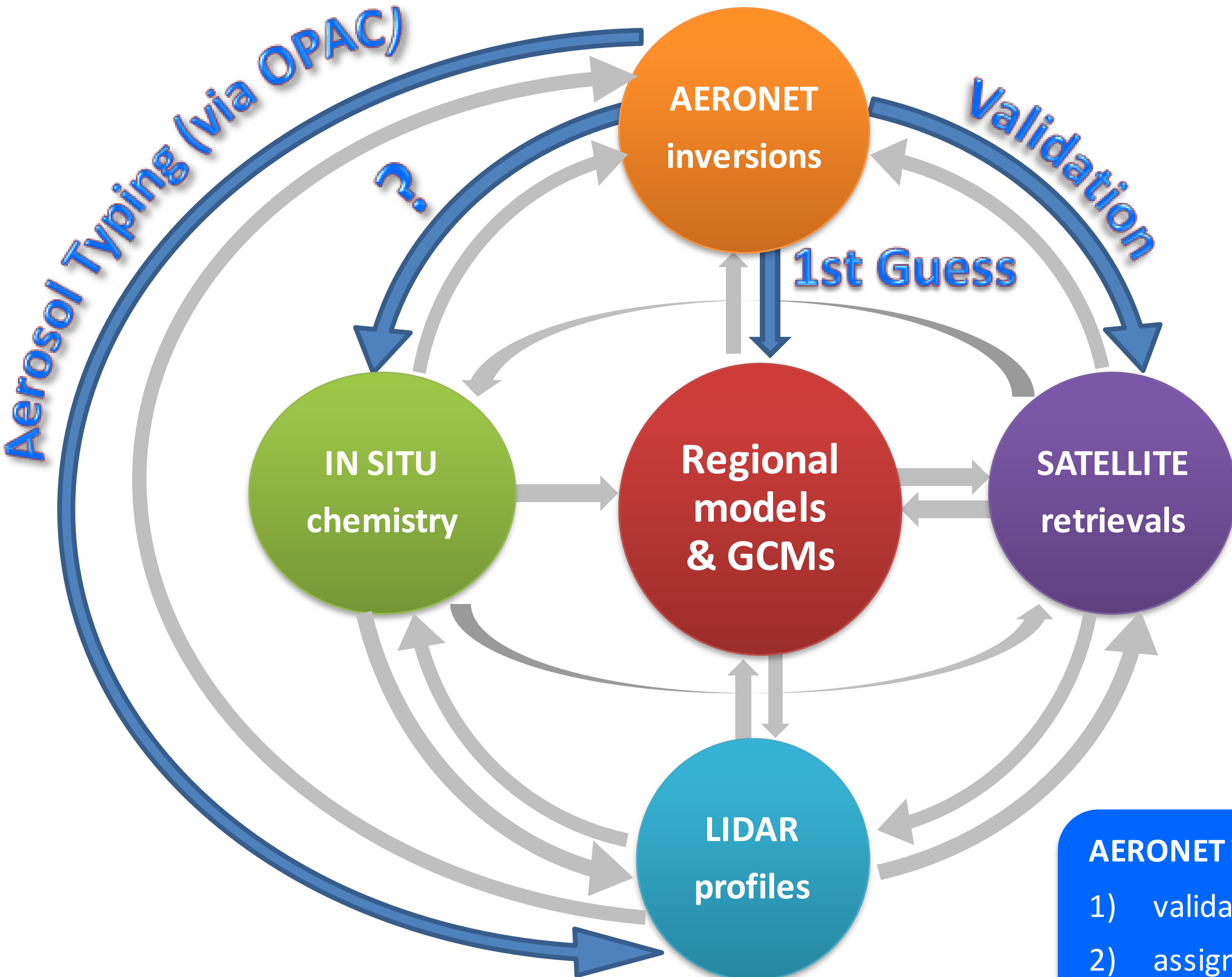
Almost **NO** change in composition but
a **LARGE change in H2O** (x2)

3a) Fresno 8th -14th Feb 2006 → *Fog-induced modification*



4. Potential impacts

4b) Potential impact: *on AERONET & knock-on effects?*

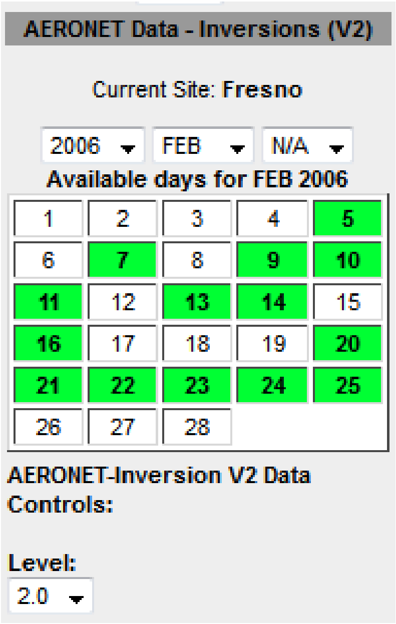


AERONET is a reference for:

1) validating satellite retrievals
2) assigning aerosol types
3) providing 1st guesses in "spin-ups"

4) A wish list

1) **More continuity in the AERONET inversion data record → to enable studies of the temporal evolution of atypical aerosol events**
2) **Establishment of a taxonomy database → to help detect, assess and monitor atypical events**
3) **Incorporation of our algorithm into operational algorithms & AERONET (inversion) data products**
4) **Your suggestions ☺**



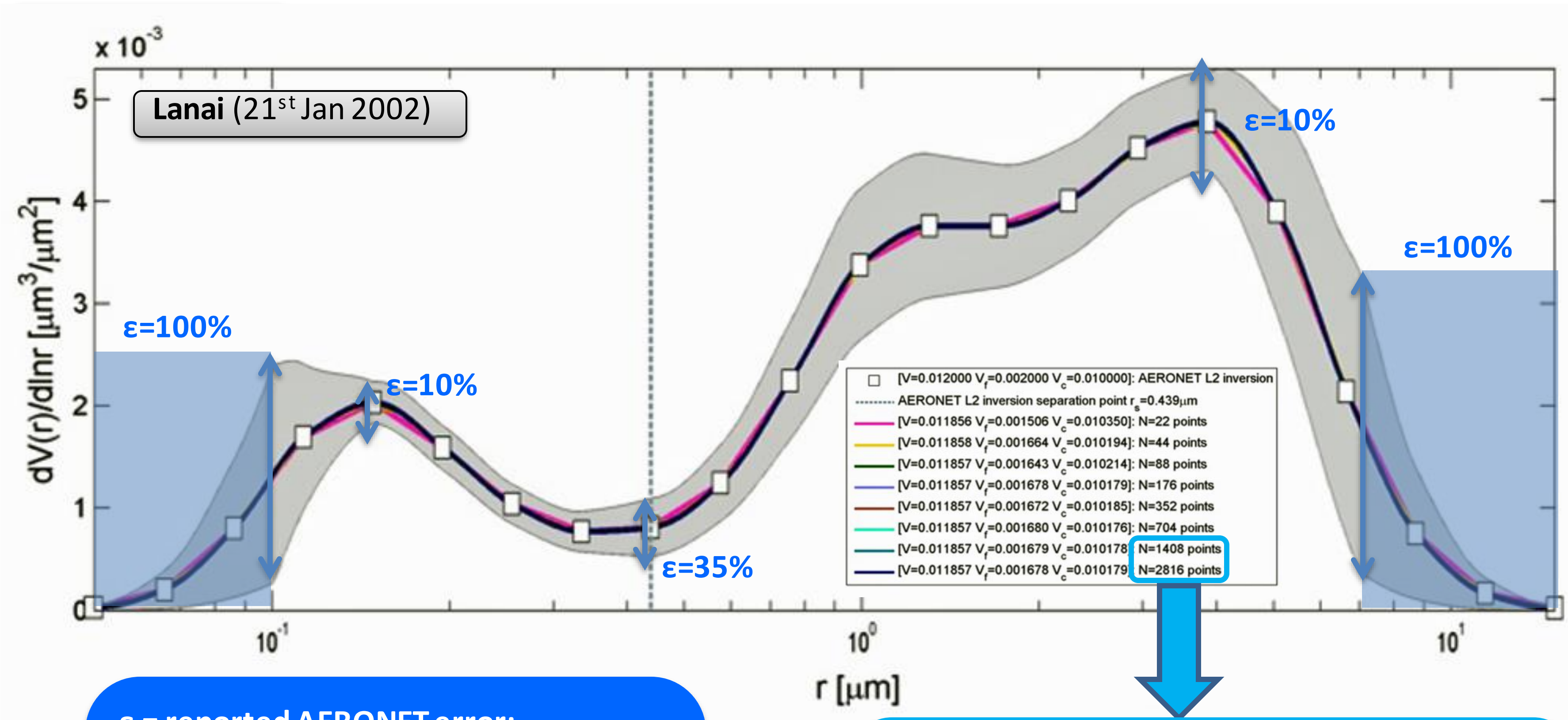
Many thanks to all our colleagues

Michael Taylor, IERSD-NOA

mtaylor@noa.gr

EXTRA SLIDES

Errors on the AERONET retrieval



ε = reported AERONET error:

ε = 10% at fine & coarse peaks
ε = 35% at local minimum
ε = 100% at edges: r < 0.1μm or r > 7μm

Dubovik et al (2002)

***N*=2200 (x10 AERONET) gave best results for:**

1) calculation of max(R^2) in the OEV method
2) stabilization of the SE in the GMM method

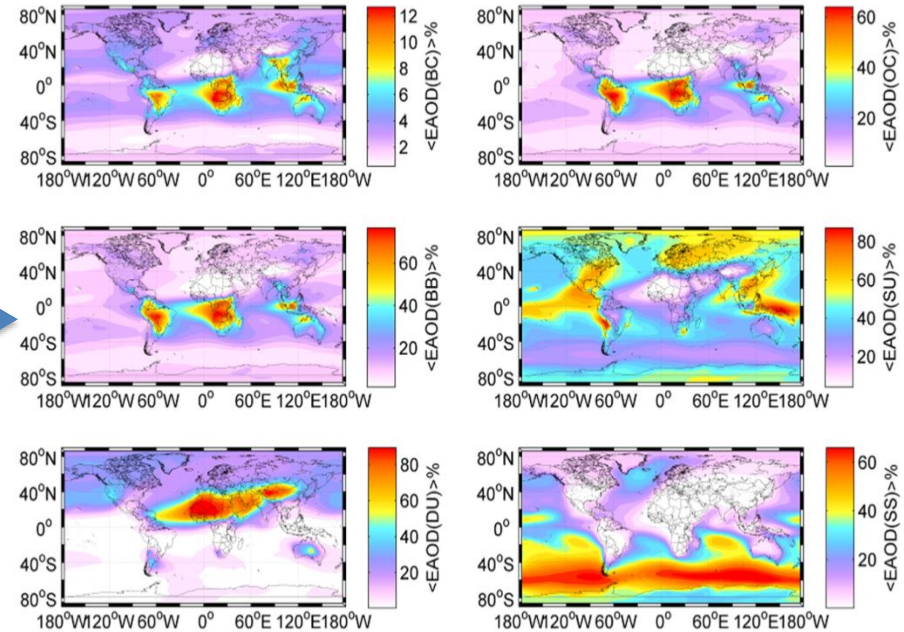
Using GOCART to isolate events

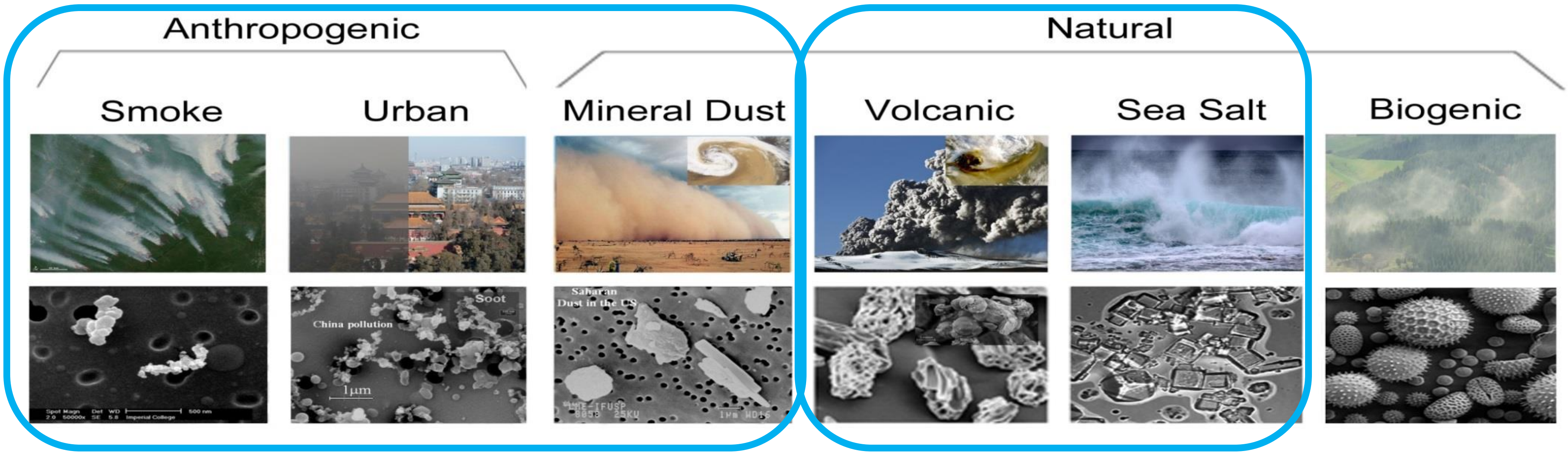
Year 2000 Emissions Tg yr⁻¹ or TgS yr⁻¹	Anthropogenic NMVOCs		Anthropogenic Black Carbon		Anthropogenic POA		Anthropogenic SO_2		Anthropogenic NH_3		Biomass Burning Aerosols	
	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
Total	98.2	157.9	3.6	6.0	6.3	15.3	43.3	77.9	34.5	49.6	29.0	85.3

Source	Natural Global	
	Min	Max
Sea spray	1400	6800
Mineral dust	1000	4000
Terrestrial PBAPs	50	1000
Dimethylsulphide (DMS)	10	40
Monoterpenes	30	120
Isoprene	410	600
SOA production from all BVOCs	20	380

Source: IPCC/AR5 (2013)

GOCART 2000-2006 global mean AOD per type





OEV method: *"pure" aerosol cases*

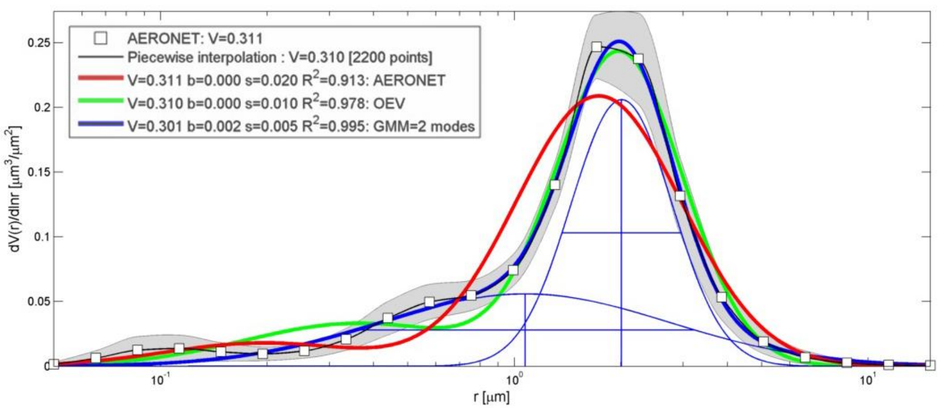
Aerosol type	AERONET site	Peak date	SU	DU	SS	OC+BC
Dust	Banizoumbou	16/03/2005	1.02%	97.91%	0.03%	1.04%
Biomass Burning	Mongu	14/08/2003	5.61%	0.22%	0.05%	94.12%
Urban SO_2	GSFC-Washington	17/08/2005	87.53%	1.38%	0.05%	11.04%
Marine (sea salt)	Lanai	21/02/2002	28.92%	3.32%	60.14%	7.61%

Aerosol type	p [r_f]	p [r_c]	p [σ_f]	p [σ_c]	p [Vf]	p [V_c]
Dust	0.014 **	0.039	0.456	0.083	0.013 **	0.013 **
Biomass Burning	0.160	0.930	0.092	0.178	0.654	0.678
Urban SO_2	0.572	0.982	0.237	0.120	0.139	0.152
Marine (sea salt)	0.017 **	0.035	0.048	0.132	0.012 **	0.008 **

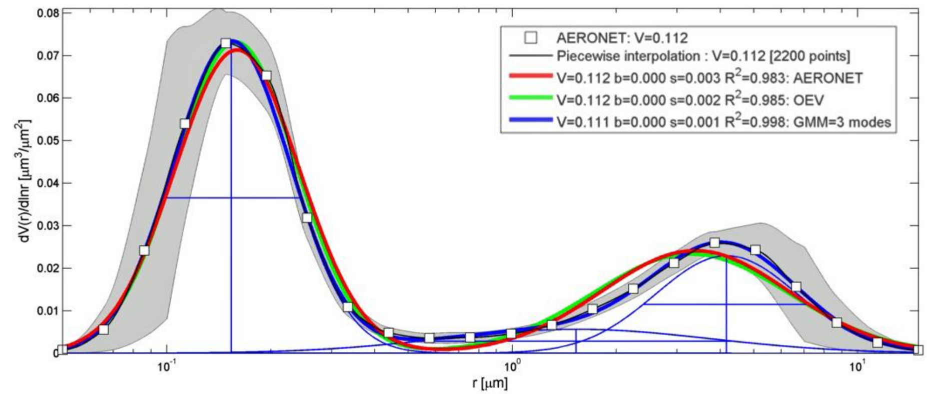
** → statistically-significant for dust & marine-dominated AVSD (2-tailed paired t-test at the 95% level of confidence: $p<0.025$)

GMM V OEV V AERONET: *"pure" aerosol cases*

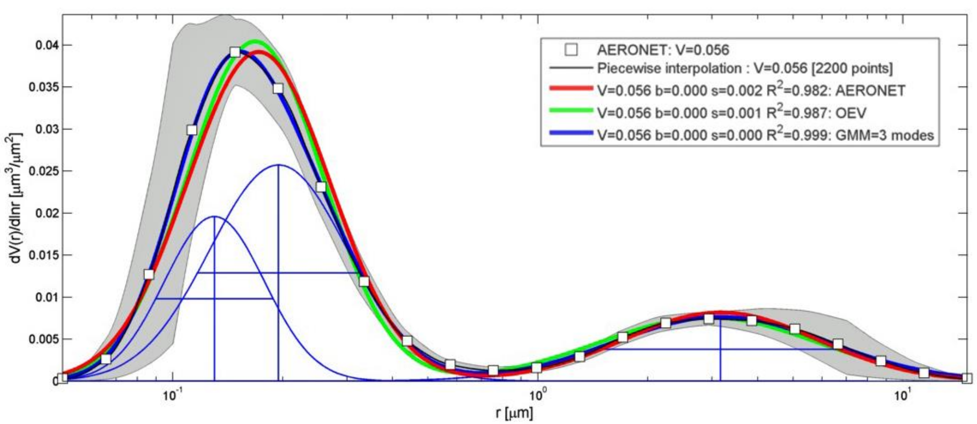
DUST (98%): Banizoumbou (16th March 2005) = "***Quenched fine mode***"



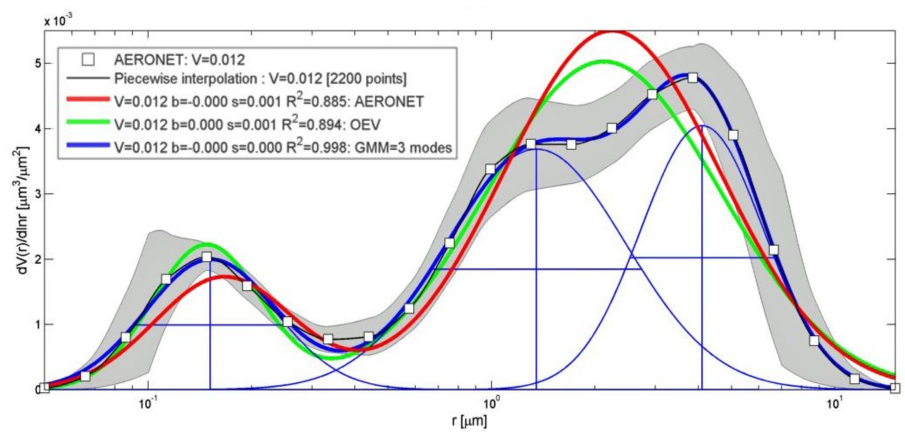
SMOKE (94%): Mongu (14th August 2003) = "***Typical***"



URBAN SO_2 (88%) : GFSC-Washington (17th August 2005) = "***Typical***"



SEA SALT (60%): Lanai (21st Jan 2002) = "***Double coarse peak***"



GMM method: *some more maths*

$$f_i = a_i e^{-\left(\frac{\ln r - b_i}{c_i}\right)^2}$$

$$\frac{dV(r)}{d\ln r} = \sum_{i=1..n} a_i e^{-\left(\frac{\ln r - b_i}{c_i}\right)^2}$$

$$\frac{\partial f_i}{\partial a_i} = a_i e^{-\left(\frac{\ln r - b_i}{c_i}\right)^2}$$

$$\frac{\partial f_i}{\partial b_i} = \frac{2a_i(\ln r - b_i)}{c_i^2} e^{-\left(\frac{\ln r - b_i}{c_i}\right)^2}$$

$$\frac{\partial f_i}{\partial c_i} = \frac{2a_i(\ln r - b_i)^2}{c_i^3} e^{-\left(\frac{\ln r - b_i}{c_i}\right)^2}$$

Derivatives are sensitive to *N*

$$V_i = \sqrt{\pi}(a_i c_i)$$

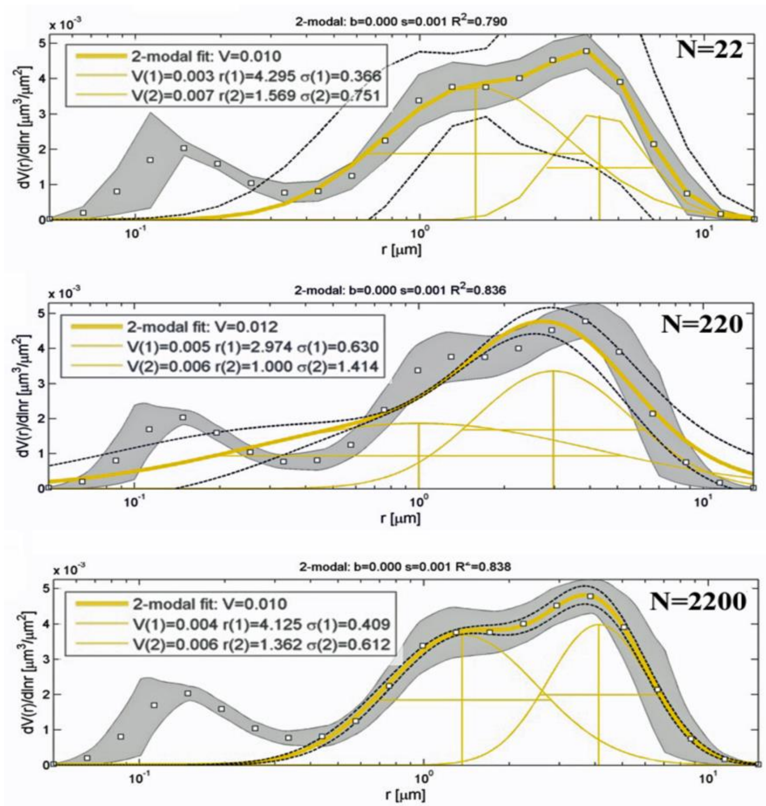
$$r_i = e^{b_i}$$

$$\sigma_i = \frac{1}{\sqrt{2}} c_i$$

$$SE_i = \sqrt{\left(\frac{\partial f_i}{\partial a_i} \times SE(a_i)\right)^2 + \left(\frac{\partial f_i}{\partial b_i} \times SE(b_i)\right)^2 + \left(\frac{\partial f_i}{\partial c_i} \times SE(c_i)\right)^2}$$

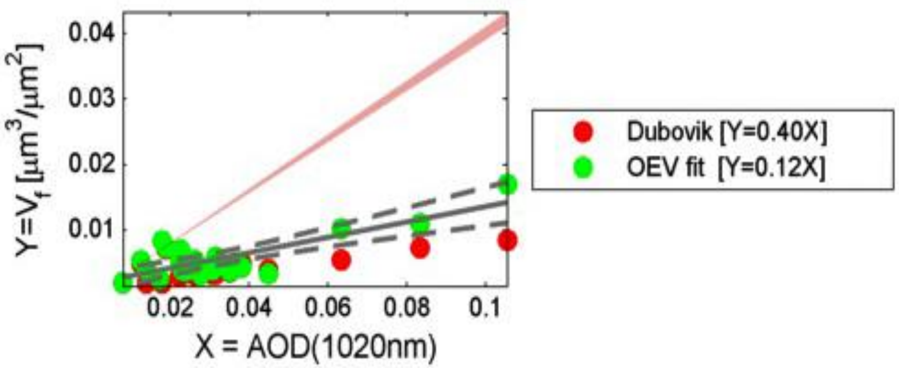
$$upper\ bound = \sum_{i=1..n} a_i e^{-\left(\frac{\ln r - b_i}{c_i}\right)^2} + 1.96\sqrt{\sum_{i=1..n} (SE_i)^2}$$

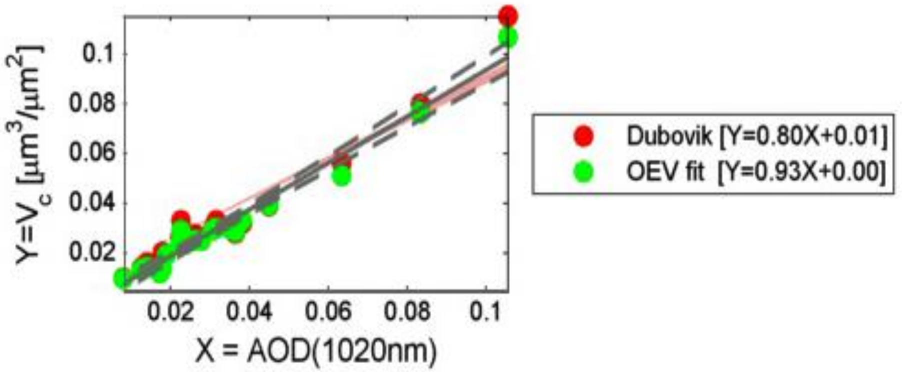
$$lower\ bound = \sum_{i=1..n} a_i e^{-\left(\frac{\ln r - b_i}{c_i}\right)^2} - 1.96\sqrt{\sum_{i=1..n} (SE_i)^2}$$



Potential impact: *on AERONET optical properties*

Lanai (Jan 2002)

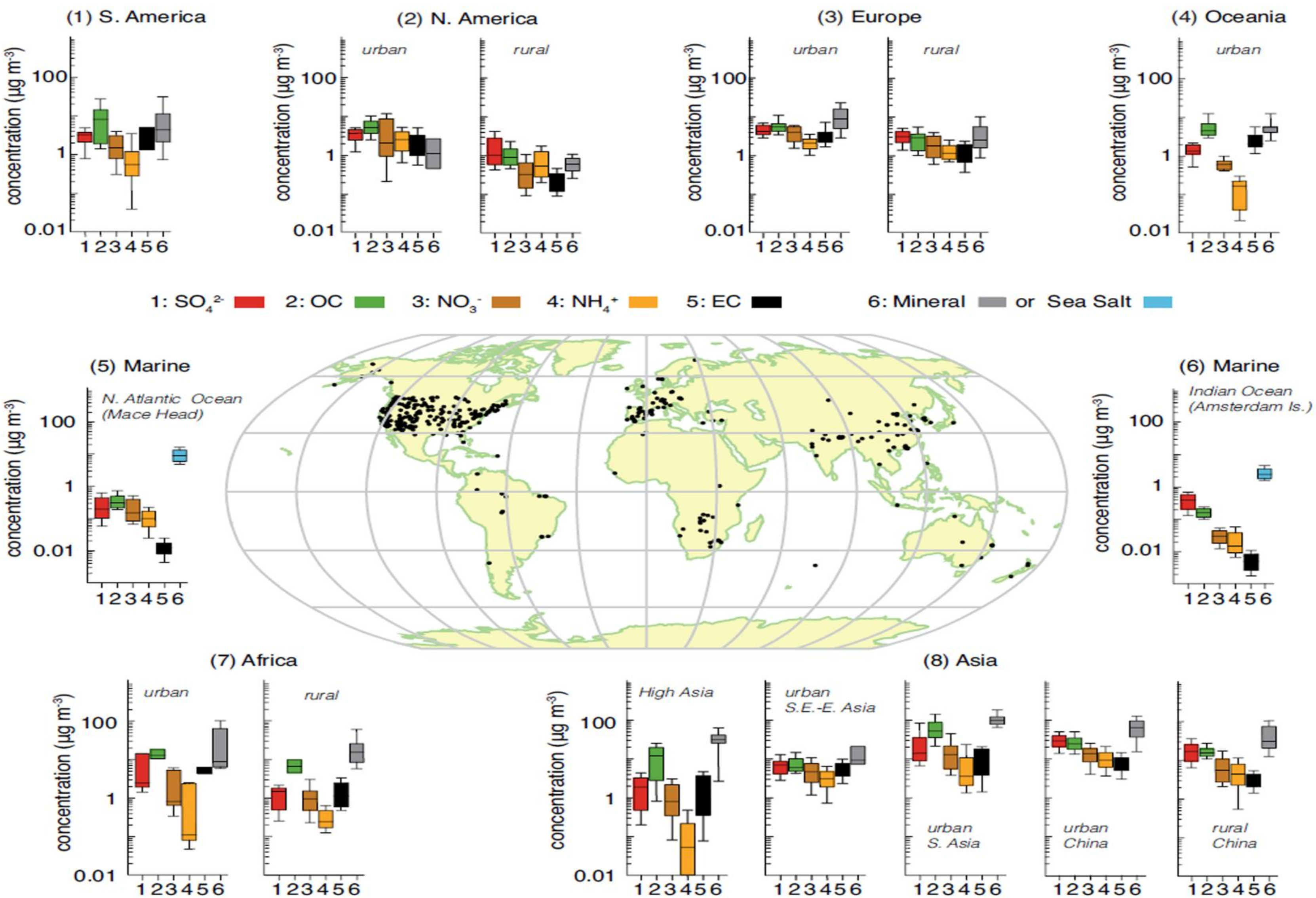




V_f = 0.12 x AOD(1020) → AOD(1020) ≈ 8.33 V_f
V_c = 0.93 x AOD(1020) → AOD(1020) ≈ 1.08 V_c
r_s(AERONET) = 0.439μm but r_s(OEV) = 0.587μm
→ RE(AERONET-OEV) ≈ -28% for V_f and ≈ +7% for V_c

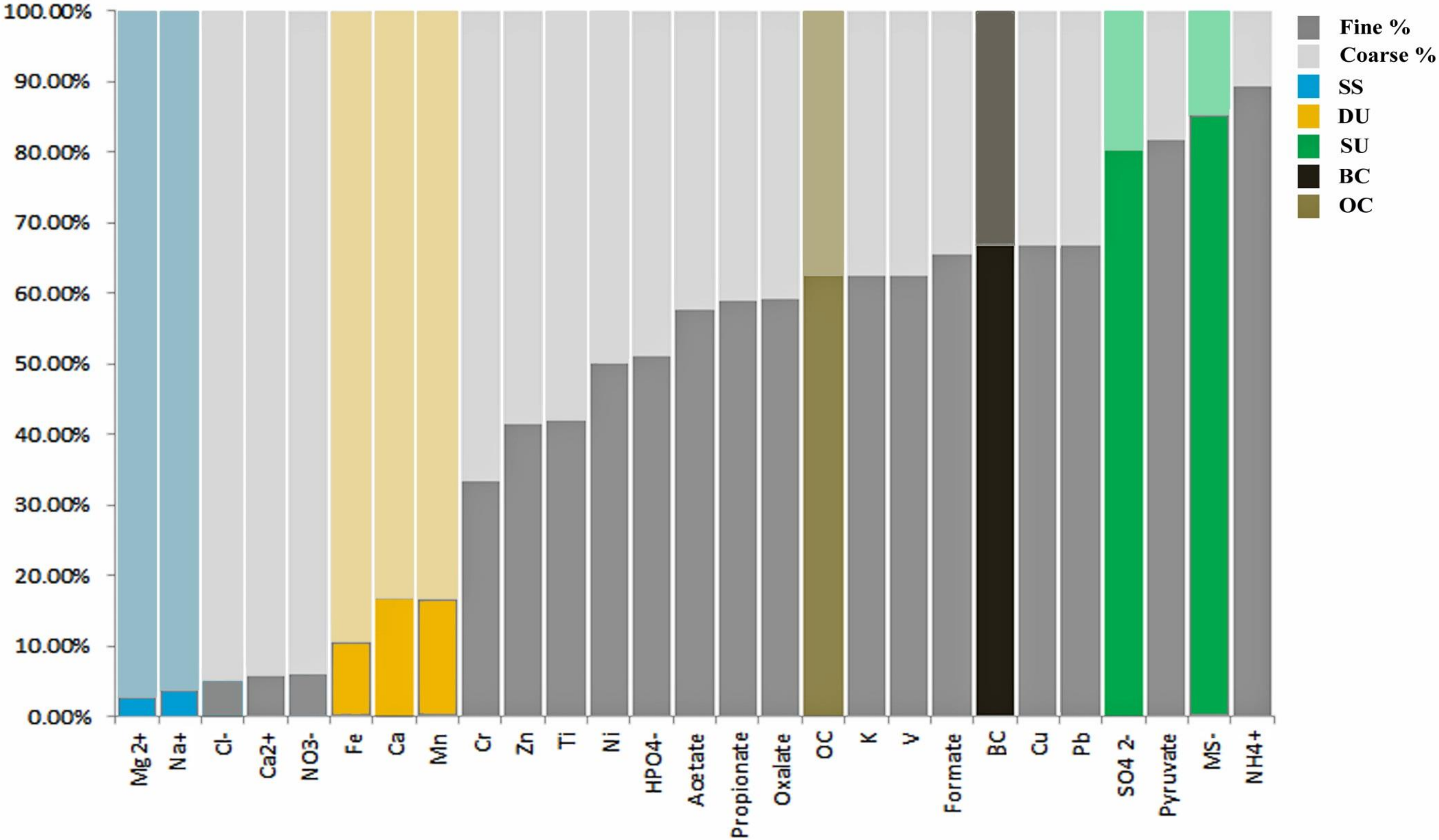
→ AOD(1020) for the fine mode ≈ 28% higher with OEV than AERONET
→ AOD(1020) for the coarse mode ≈ 7% lower with OEV than AERONET

Global PM10 concentrations



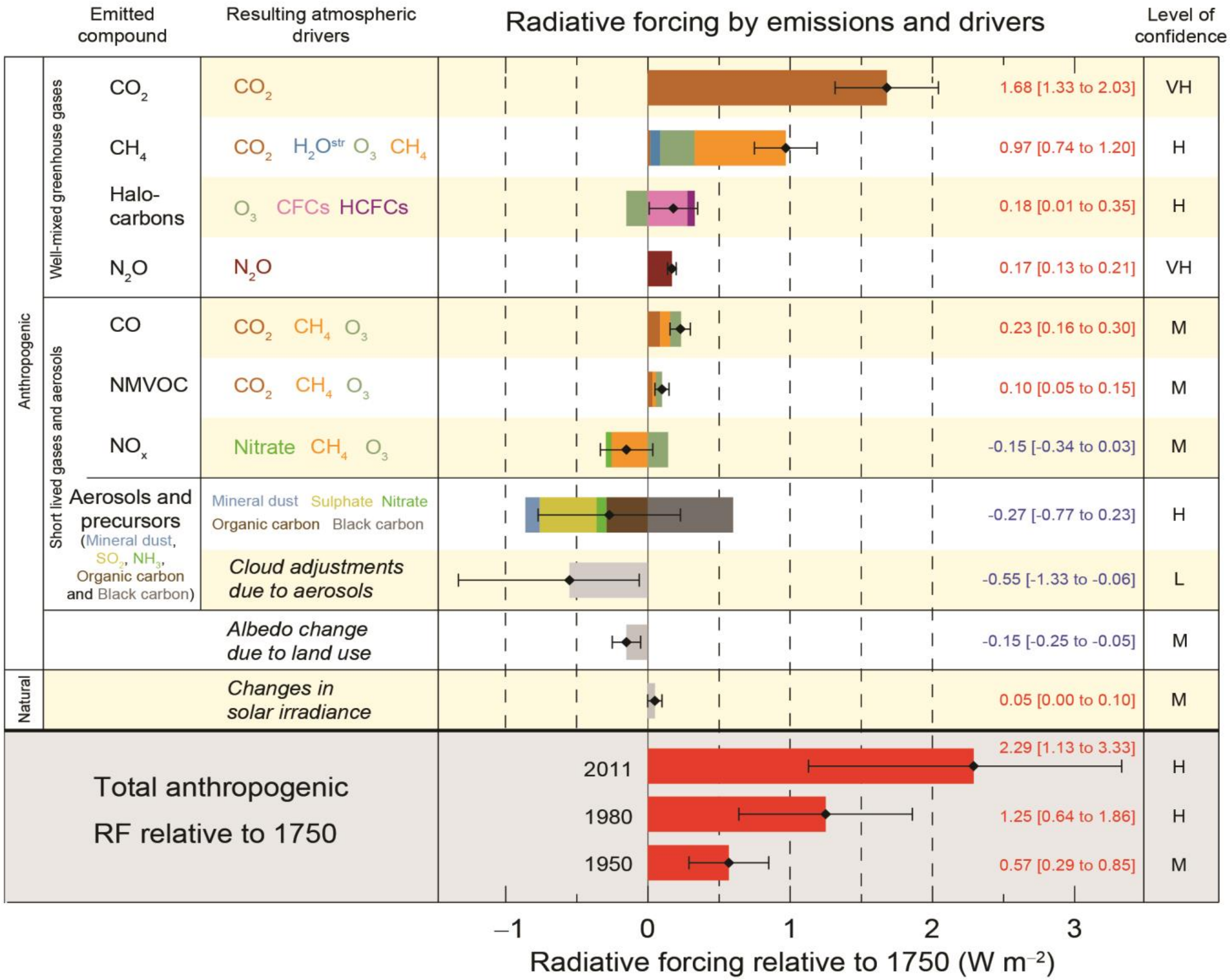
Source: IPCC/AR5 (2013)

Fine & coarse mode chemistry



After: Gerasopoulos et al (2007)

Aerosol -radiation-cloud uncertainty is confidently LARGE



***Source**: IPCC/AR5 (2013)*

Accurate V Precise

Tropospheric aerosol lifetimes

Aerosol Species	Size Distribution	Main Sources	Main Sinks	Tropospheric Lifetime	Key Climate Relevant Properties
Sulphate	Primary: Aitken, accumulation and coarse modes Secondary: Nucleation, Aitken, and accumulation modes	Primary: marine and volcanic emissions. Secondary: oxidation of SO_2 and other S gases from natural and anthropogenic sources	Wet deposition Dry deposition	~ 1 week	Light scattering. Very hygroscopic. Enhances absorption when deposited as a coating on black carbon. Cloud condensation nuclei (CCN) active.
Nitrate	Accumulation and coarse modes	Oxidation of NO_x	Wet deposition Dry deposition	~ 1 week	Light scattering. Hygroscopic. CCN active.
Black carbon	Freshly emitted: <100 nm Aged: accumulation mode	Combustion of fossil fuels, biofuels and biomass	Wet deposition Dry deposition	1 week to 10 days	Large mass absorption efficiency in the shortwave. CCN active when coated. May be ice nuclei (IN) active.
Organic aerosol	POA: Aitken and accumulation modes. SOA: nucleation, Aitken and mostly accumulation modes. Aged OA: accumulation mode	Combustion of fossil fuel, biofuel and biomass. Continental and marine ecosystems. Some anthropogenic and biogenic non-combustion sources	Wet deposition Dry deposition	~ 1 week	Light scattering. Enhances absorption when deposited as a coating on black carbon. CCN active (depending on aging time and size).
Mineral dust	Coarse and super-coarse modes, with a small accumulation mode	Wind erosion, soil resuspension. Some agricultural practices and industrial activities (cement)	Sedimentation Dry deposition Wet deposition	1 day to 1 week depending on size	IN active. Light scattering and absorption. Greenhouse effect.
Sea spray	Coarse and accumulation modes	Breaking of air bubbles induced e.g., by wave breaking. Wind erosion.	Sedimentation Wet deposition Dry deposition	1 day to 1 week depending on size	Light scattering. Very hygroscopic. CCN active. Can include primary organic compounds in smaller size range

***Source**:* IPCC/AR5 (2013)